

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 9, 2003, 16:11:58 ; Search time 9.77143 Seconds
(without alignments)
91.441 Million cell updates/sec

Title: US-09-905-691-3
Perfect score: 19
Sequence: 1 AEARARAAARAAARA 19

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 127863 seqs, 47026705 residues

Word size : 0

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	9	47.4	756	1 NRG2_MOUSE	P56974 mus musculus
2	9	47.4	850	1 NRG2_HUMAN	O14511 homo sapien
3	9	47.4	868	1 NRG2_RAT	O35569 rattus norv
4	8	42.1	448	1 TRME_XANAC	O8peh9 xanthomonas
5	8	42.1	2252	1 POLI_GCMV	P13025 h rnal poly
6	7	36.8	126	1 TYRT_STRAL	P55046 streptomyc
7	7	36.8	134	1 TYRT_STRLA	P55047 streptomyc
8	7	36.8	140	1 TYRT_STRLN	P55048 streptomyc
9	7	36.8	145	1 YV40_DEIRA	O9zrs4 deinococcus
10	7	36.8	146	1 TYRT_STRA	P17687 streptomyc
11	7	36.8	193	1 VCO7_ADE04	O96831 human adeno
12	7	36.8	219	1 DEDA_ECOLI	O9548 escherichia
13	7	36.8	256	1 TRUA_METKA	O8twz3 methanopyru
14	7	36.8	323	1 YOC5_MICLE	O69538 mycobacteri
15	7	36.8	326	1 YOC5_MICTU	P37895 caulobacteri
16	7	36.8	344	1 HEM2_MYCTU	P71765 mycobacteri
17	7	36.8	362	1 CKRA_HUMAN	P46092 homo sapien
18	7	36.8	363	1 VDH_STRAL	O69056 streptomyc
19	7	36.8	381	1 ARGJ_THERH	P96137 thermus the
20	7	36.8	383	1 SUC2_STRCO	O87840 streptomyc
21	7	36.8	387	1 ARGJ_METKA	O8tx15 m arginine
22	7	36.8	391	1 DXR_RHIME	O921p6 rhizobium m
23	7	36.8	394	1 FXD3_CHICK	P79772 gallus gall
24	7	36.8	444	1 YOR3_GLUOX	O05543 gluconobact
25	7	36.8	446	1 TRME_XANCP	O8p340 xanthomonas
26	7	36.8	451	1 TRME_XYLFA	O9p9u3 xylella fas
27	7	36.8	499	1 GSHR_PLAF7	O15770 plasmodium
28	7	36.8	499	1 GSHR_PLAFK	O94655 plasmodium
29	7	36.8	653	1 CIK4_HUMAN	P22459 homo sapien
30	7	36.8	735	1 TREA_SCHPO	O42893 schizosacch
31	7	36.8	756	1 METE_BACHD	O9kfp1 bacillus ha
32	7	36.8	799	1 SYFB_CAUCR	O9a9e5 caulobacter
33	7	36.8	859	1 ENV_HV2CA	P24105 human immun

34	7	36.8	974	1 PODJ_CAUCR	O9zg88 caulobacter
35	6	31.6	68	1 VHTV_SCOL6	P03727 escherichia
36	6	31.6	74	1 YOR2_BFP2	O06424 bacterioph
37	6	31.6	80	1 YNO2_MYCTU	O50663 mycobacteri
38	6	31.6	94	1 R28A_MYCTU	Q10879 mycobacteri
39	6	31.6	117	1 R28A_HAEIN	P44356 haemophilus
40	6	31.6	117	1 VGLJ_HSVSB	P36344 simian heip
41	6	31.6	134	1 YF61_MYCTU	Q10770 mycobacteri
42	6	31.6	135	1 NIUL_RHOCA	Q07178 rhodobacter
43	6	31.6	136	1 BLS_STRSJ	P19997 streptococc
44	6	31.6	137	1 RS11_METKA	Q8tvb9 methanopyru
45	6	31.6	137	1 RS11_PYRFU	Q8u0e3 pyrococcus

ALIGNMENTS

RESULT 1
NRG2_MOUSE
ID NRG2_MOUSE STANDARD; PRT; 756 AA.
AC P56974;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Pro-neuregulin-2 precursor (Pro-NRG2) [Contains: Neuregulin-2 (NRG-2)
DE (Divergent of neuregulin 1) (DON-1)].
GN NRG2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS NRG2-5; NRG2-10 AND NRG2-16A).
RC STRAIN=C57BL/6; TISSUE=Brain;
RX MEDLINE=97311398; PubMed=9168115;
RA Carraway K.L. III, Weber J.L., Unger M.J., Ledesma J., Yu N.,
RA Gassmann M., Lai C.;
RT "Neuregulin-2, a new ligand of ErbB3/ErbB4-receptor tyrosine
RT kinases".
RL Nature 387:512-516(1997).
RN [2]
RP SEQUENCE OF 150-756 FROM N.A. (ISOFORMS DON-1M AND DON-1S).
RC TISSUE=Choroid plexus;
RX MEDLINE=97342638; PubMed=9199335;
RA Busfield S.J., Michnick D.A., Chickering T.W., Revett T.L., Ma J.,
RA Woolf E.A., Comrack C.A., Dussault B.J., Woolf J., Goodearl A.D.J.,
RA Gearing D.P.;
RT "Characterization of a neuregulin-related gene, Don-1, that is highly
RT expressed in restricted regions of the cerebellum and hippocampus".
RL Mol. Cell. Biol. 17:4007-4014(1997).
CC -1- FUNCTION: DIRECT LIGAND FOR ERBB3 AND ERBB4 TYROSINE KINASE
CC RECEPTORS. CONCOMITANTLY RECRUITS ERBB1 AND ERBB2 CORECEPTORS,
CC RESULTING IN LIGAND-STIMULATED TYROSINE PHOSPHORYLATION AND
CC ACTIVATION OF THE ERBB RECEPTORS. MAY ALSO PROMOTE THE
CC HETERODIMERIZATION WITH THE EGF RECEPTOR.
CC -1- SUBCELLULAR LOCATION: EXISTS AS AN TYPE I MEMBRANE PROTEIN AND AS
CC A PROTEOLYTICALLY RELEASED SOLUBLE GROWTH FACTOR FORM. THE
CC MEMBRANE-BOUND FORM DOES NOT SEEM TO BE ACTIVE (BY SIMILARITY).
CC -1- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing: Named isoforms-4;
CC Comment-Additional isoforms seem to exist;
CC Name-NRG2-16A;
CC IsoId=P56974-1; Sequence-Displayed;
CC Name-DON-1M;
CC IsoId=P56974-2; Sequence-VSP_003464;
CC Name-DON-1S; Synonyms-NRG2-5;
CC IsoId=P56974-3; Sequence-VSP_003462, VSP_003463;
CC Name-NRG2-10;
CC IsoId=P56974-4; Sequence-VSP_003460, VSP_003461;
CC -1- TISSUE SPECIFICITY: HIGHEST EXPRESSION IN THE BRAIN, WITH LOWER
CC LEVELS IN THE LUNG. IN THE CEREBELLUM, FOUND IN GRANULE AND
CC PURKINJE CELLS.
CC -1- DOMAIN: THE CYTOPLASMIC DOMAIN MAY BE INVOLVED IN THE REGULATION


```

FT VARSPLIC 414 421 Missing (in isoform 2 and isoform 3).
FT FTID-VSP_003471.
FT HIGFELKEAEELYQKRVLTIGICVA -> SVLWDPGPGV
FT SSSONSTSPSTLDLN (in isoform 6).
FT FTID-VSP_003472.
FT Missing (in isoform 6).
FT FTID-VSP_003473.
FT CONFLICT 117 117 S -> F (IN REF. 2).
FT CONFLICT 724 724 R -> H (IN REF. 2).
SQ SEQUENCE 868 AA; 93776 MW; 3C7D4D94DBE64DE2 CRC64;

Query Match 47.4%; Score 9; DB 1; Length 868;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 EEARARRAAA 10
DB 578 EEARARRAAA 586
|||||

RESULT 4
TRME_XANAC STANDARD; PRT; 448 AA.
AC Q8PEH9;
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE tRNA modification GTPase trme.
GN TRME OR THDF OR YAC4370.
OS Xanthomonas axonopodis (pv. citri).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OC NCBI_TaxID=92829;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=306 / ATCC 13902 / XV 101;
RX MEDLINE=22022145; PubMed=12024217;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA Camarotte G., Cannavan F., Cardozo J., Chambergo F., Clapina L.P.,
RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities."
RL Nature 417:459-463(2002).
CC -1- FUNCTION: Exhibits a very high intrinsic GTPase hydrolysis rate.
CC Involved in the biosynthesis of the hypermodified nucleoside 5-
CC methylaminomethyl-2-thiouridine, which is found in the wobble
CC position of some tRNAs (By similarity).
CC -1- SIMILARITY: Belongs to the era/trme family of GTP-binding
CC proteins. Trme subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE012091; AM39200.1; -
CC HAMAP; MF_00379; -; 1.
CC InterPro; IPR005289; GTP-binding_dom.

```

```

DR InterPro; IPR006073; GTP1_OBG.
DR InterPro; IPR005225; Small_GTP.
DR InterPro; IPR004520; ThdF.
DR PRINTS; PRO0326; GTP1_OBG.
DR TIGRFAMS; TIGR00650; MG442; 1.
DR TIGRFAMS; TIGR00231; small_GTP; 1.
DR TIGRFAMS; TIGR00450; thdF; 1.
KW tRNA processing; GTP-binding; Complete proteome.
FT NP_BIND 225 232 GTP (POTENTIAL).
FT NP_BIND 272 276 GTP (POTENTIAL).
FT NP_BIND 332 335 GTP (POTENTIAL).
SQ SEQUENCE 448 AA; 47543 MW; B6B43D163D92E3F4 CRC64;

Query Match 42.1%; Score 8; DB 1; Length 448;
Best Local Similarity 100.0%; Pred. No. 4.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 AARAARRA 16
DB 142 AARAARRA 149
|||||

RESULT 5
POL1_GCMV STANDARD; PRT; 2352 AA.
ID POL1_GCMV
AC P13025;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE RNAI polypeptide (250 kDa protein) [Contains: 63 kDa protease
DE cofactor; 72 kDa membrane-binding protein; Genome-linked protein
DE (VPG); Protease (EC 3.4.22.-); RNA-directed RNA polymerase
DE (EC 2.7.7.48)].
DE Hungarian grapevine chrome mosaic virus (GCMV).
OS Viruses; ssRNA positive-strand viruses, no DNA stage; Comoviridae;
OC Nepovirus.
OX NCBI_TaxID=12273;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9001864; PubMed=2798128;
RA le Gall O., Candresse T., Braut V., Dunez J.;
RT "Nucleotide sequence of Hungarian grapevine chrome mosaic nepovirus
RT RNAI."
RL Nucleic Acids Res. 17:7795-7807(1989).
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
CC [RNA](N).
CC -1- PTM: CLEAVAGE OF ONE OF THE R/A DIPEPTIDES AT POSITIONS 249, 252
CC & 259 COULD YIELD THE PUTATIVE N-TERMINAL PROTEIN OF NEPOVIRUSES.
CC -1- SIMILARITY: WITH RNA-1 POLYPROTEIN OF TOMATO BLACK RING VIRUS.
CC (TBV) (638) AND WITH POLYPROTEIN B OF COMPEA MOSAIC VIRUS (CPMV)
CC (218).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X15346; CAA33405.1; -
CC PIR; S06188; S06188.
CC MEROPS; C03.004; -
CC InterPro; IPR004004; Calici_pol_hel.
CC InterPro; IPR000605; RNA_helicase.
CC InterPro; IPR007095; RNA_pol_DS_PS.
CC InterPro; IPR001205; RNA_pol_P3D.
CC InterPro; IPR007094; RNA_pol_PSVir.
CC Pfam; PF00680; RNA_dep_RNA_pol; 1.
CC Pfam; PF00910; RNA_helicase; 1.
CC PRINTS; PR00918; CALICIVIRUSNS.
CC PolyProtein; Transmembrane; Hydrolase; Protease; Transferase;
CC RNA-directed RNA polymerase; ATP-binding.

```

```

FT CHAIN 1 ? 63 kDa PROTEASE COFACTOR (POTENTIAL).
FT CHAIN ? ? 72 kDa MEMBRANE-BINDING PROTEIN
FT CHAIN ? ? (POTENTIAL).
FT CHAIN ? ? GENOME-LINKED PROTEIN (POTENTIAL).
FT CHAIN ? 1428 CYSTEINE PROTEASE (POTENTIAL).
FT CHAIN 1429 2252 RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
FT NP_BIND 763 770 ATP (POTENTIAL).
FT ACT_SITE 1386 1386 CYSTEINE PROTEASE (POTENTIAL).
SQ SEQUENCE 2252 AA: 249865 MW; DDFB0DCF7E94171B CRC64;

Query Match 42.1%; Score 8; DB 1; Length 2252;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 4 RAAARAA 11
Db 249 RAAARAA 256

RESULT 6
TYRT_STRAL STANDARD; PRT; 126 AA.
AC P55046;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Tyrosinase co-factor.
GN MELC1.
OS Streptomyces albus G.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1962;
RN [1]
RC STRAIN=DSM 40480;
RA Weimeler U.F., Brass N., Roessler C., Piepersberg W.;
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: THIS PROTEIN MAY FUNCTION TO DELIVER COPPER TO
CC TYROSINASE.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL; X95705; CAA65004.1; -
CC Melanin biosynthesis; Copper.
KW SEQUENCE 126 AA; 12916 MW; 1785CC2C777C0106 CRC64;

Query Match 36.8%; Score 7; DB 1; Length 126;
Best Local Similarity 100.0%; Pred. No. 9.7;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 7 RAAARAA 13
Db 105 RAAARAA 111

RESULT 7
TYRT_STRGA STANDARD; PRT; 134 AA.
AC P55047;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Tyrosinase co-factor (URF402).
GN MELC1.
OS Streptomyces glaucescens.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.

```

```

OX NCBI_TaxID=1907;
RN [1]
RC SEQUENCE FROM N.A.
RC STRAIN=ATCC 13032 / ETH 22794 / GLA.0;
RX MEDLINE=88040431; PubMed=3118334;
RA Huber M., Huetter R., Lerch K.;
RT "The promoter of the Streptomyces glaucescens mel operon.";
RL Nucleic Acids Res. 15:8106-8106(1987).
CC -1- FUNCTION: THIS PROTEIN MAY FUNCTION TO DELIVER COPPER TO
CC TYROSINASE.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL; Y00457; CAA68512.1; -
CC PIR; A26986; A26986.
CC InterPro; IPR006311; Tat.
CC TIGRFAMs; TIGR01409; Tat_signal_seq; 1.
KW Melanin biosynthesis; Copper.
SQ SEQUENCE 134 AA; 13593 MW; 1E874145A7D1B9AE CRC64;

Query Match 36.8%; Score 7; DB 1; Length 134;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 7 RAAARAA 13
Db 113 RAAARAA 119

RESULT 8
TYRT_STRLN STANDARD; PRT; 140 AA.
AC P55048;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Tyrosinase co-factor.
GN MELC1.
OS Streptomyces lincolnensis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1915;
RN [1]
RC SEQUENCE FROM N.A.
RC STRAIN=78-11;
RA Zhang H.Z., Piepersberg W.;
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: THIS PROTEIN MAY FUNCTION TO DELIVER COPPER TO
CC TYROSINASE.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL; X95703; CAA64999.1; -
CC Melanin biosynthesis; Copper.
KW SEQUENCE 140 AA; 14189 MW; 8752156617FD15A6 CRC64;

Query Match 36.8%; Score 7; DB 1; Length 140;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 7 RAAARAA 13

```

```

Db          119 RAAARAA 125
|||||
RESULT 9
YV40_DEIRA STANDARD; PRT; 145 AA.
AC Q9RZS4;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE Hypothetical protein DB0040.
GN DB0040.
OS Deinococcus radiodurans.
OC Plasmid Mpl.
CC Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;
CC Deinococcaceae; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=R1 / ATCC 13939 / DSM 20539 / NCIB 9279;
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zalewski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT "Genome sequence of the radioresistant bacterium Deinococcus
RT radiodurans R1.";
RL Science 286:1571-1577(1999).
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; AF001826; AAF12646.1; -
DR PIR; E75622; E75622.
DR TIGR; DRB0040; -.
DR InterPro; IPR001844; Chaprin_Cpn60.
KW Hypothetical protein; Plasmid; Complete proteome.
SQ SEQUENCE 145 AA; 15152 MW; B8ED524495897EC3 CRC64;
-----
Query Match 36.8%; Score 7; DB 1; Length 145;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 AARRAR 18
Db 132 AARRAR 138
|||||
RESULT 10
TYRT_STRAT STANDARD; PRT; 146 AA.
AC P17687;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Tyrosinase co-factor (Orf438).
GN MELC1.
OS Streptomyces antibioticus.
CC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
CC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1890;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86031341; PubMed=3932128;
-----
Query Match 36.8%; Score 7; DB 1; Length 145;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 AARRAR 18
Db 132 AARRAR 138
|||||
RESULT 11
VC07_ADE04 STANDARD; PRT; 193 AA.
AC Q96831;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Major core protein precursor (Protein VII) (pVII).
GN PVII.
OS Human adenovirus type 4.
OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
OX NCBI_TaxID=28280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Isolate RI-6;
RA Tarassishin L., Szawlowski P.W.S., McLay J., Russell W.C.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; U70921; AAC83411.1; -
DR InterPro; IPR004912; Adeno_VII.
DR Pfam; PF03228; Adeno_VII; 1.
FT PROPEP 1 24 BY SIMILARITY.
FT CHAIN 25 193 MAJOR CORE PROTEIN.
FT SITE 24 25 CLEAVAGE (BY ADENOVIRUS PROTEASE)
FT (POTENTIAL).
SQ SEQUENCE 193 AA; 21358 MW; 43137E07DB379DD0 CRC64;
-----
RA Bernan V., Filpula D., Herber W., Bibb M.J., Katz E.;
RT "The nucleotide sequence of the tyrosinase gene from Streptomyces
RT antibioticus and characterization of the gene product.";
RL Gene 37:101-110(1985).
RN [2]
RP FUNCTION.
RX MEDLINE=88284382; PubMed=2840357;
RA Lee Y.-H.W., Chen B.-F., Wu S.-Y., Leu W.-M., Lin J.-J., Chen C.C.,
RA Lo S.J.;
RT "A trans-acting gene is required for the phenotypic expression of a
RT tyrosinase gene in Streptomyces.";
RL Gene 65:71-81(1988).
CC -|- FUNCTION: THIS PROTEIN MAY FUNCTION TO DELIVER COPPER TO
CC TYROSINASE.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; M11582; AAN88570.1; -
DR PIR; A23971; A23971.
DR InterPro; IPR006311; Tat.
DR TIGRFAMs; TIGR01409; TAT_signal_seq; 1.
KW Melanin biosynthesis; Copper.
SQ SEQUENCE 146 AA; 14883 MW; E384D3CALLIAC77AF CRC64;
-----
Query Match 36.8%; Score 7; DB 1; Length 146;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 RAAARAA 13
Db 125 RAAARAA 131
|||||
RESULT 11
VC07_ADE04 STANDARD; PRT; 193 AA.
AC Q96831;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Major core protein precursor (Protein VII) (pVII).
GN PVII.
OS Human adenovirus type 4.
OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
OX NCBI_TaxID=28280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Isolate RI-6;
RA Tarassishin L., Szawlowski P.W.S., McLay J., Russell W.C.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; U70921; AAC83411.1; -
DR InterPro; IPR004912; Adeno_VII.
DR Pfam; PF03228; Adeno_VII; 1.
FT PROPEP 1 24 BY SIMILARITY.
FT CHAIN 25 193 MAJOR CORE PROTEIN.
FT SITE 24 25 CLEAVAGE (BY ADENOVIRUS PROTEASE)
FT (POTENTIAL).
SQ SEQUENCE 193 AA; 21358 MW; 43137E07DB379DD0 CRC64;
-----

```

Query Match 36.8%; Score 7; DB 1; Length 193;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RAARRAA 17
 DB 135 RAARRAA 141

RESULT 12

DEDA_ECOLI STANDARD; PRT; 219 AA.
 AC P09548;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-MAR-1989 (Rel. 10, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Deda protein (DSG-1 protein).
 GN DEDA OR B2317 OR 23579 OR ECS3201.
 OS Escherichia coli, and
 OS Escherichia coli O157:H7.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=562, 83334;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12;
 RX MEDLINE=97308226; PubMed=3040734;
 RA Nonet M.L., Marvel C.C., Tolan D.R.;
 RA "The hist-purF region of the Escherichia coli K-12 chromosome.
 RT Identification of additional genes of the hist and purF operons."
 RL J. Biol. Chem. 262:12209-12217(1987).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12 / MG1655;
 RX MEDLINE=9746617; PubMed=9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12."
 RL Science 277:1453-1474(1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12;
 RX MEDLINE=97349980; PubMed=9205837;
 RA Yamamoto Y., Alba H., Baba T., Hayashi K., Inada T., Isono K.,
 RA Itoh T., Kimura S., Kitagawa M., Makino K., Miki T., Mitsuhashi N.,
 RA Mizobuchi K., Mori H., Nakade S., Nakamura Y., Nashimoto H.,
 RA Oshima T., Oyama S., Saito N., Sampei G., Sato Y., Sivasubraman S.,
 RA Tagami H., Takahashi H., Takeda J., Takemoto K., Uehara K., Wada C.,
 RA Yamagata S., Horiuchi T.;
 RT "Construction of a contiguous 874-kb sequence of the Escherichia coli
 K-12 genome corresponding to 50.0-60.8 min on the linkage map and
 analysis of its sequence features."
 RL DNA Res. 4:91-113(1997).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN-O157:H7 / ATCC 700927;
 RX MEDLINE=21074935; PubMed=11206551;
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
 RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
 RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
 RA Welch R.A., Blattner F.R.;
 RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7."
 RL Nature 409:529-533(2001).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN-O157:H7 / RMD 0509952;
 RX MEDLINE=21156231; PubMed=11238796;
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,

RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
 RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
 RA Kuwara S., Shiba T., Hattori M., Shinagawa H.;
 RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
 O157:H7 and genomic comparison with a laboratory strain K-12."
 RL DNA Res. 8:11-22(2001).
 RN [6]
 RP SEQUENCE OF 1-9 FROM N.A.
 RC STRAIN-K12;
 RX MEDLINE=87137258; PubMed=3029016;
 RA Arps P.J., Winkler M.E.;
 RT "Structural analysis of the Escherichia coli K-12 hist operon by
 using a kanamycin resistance cassette."
 RL J. Bacteriol. 169:1061-1070(1987).
 RN [7]
 RP SEQUENCE OF 204-219 FROM N.A.
 RX MEDLINE=93123150; PubMed=7678242;
 RA Li S.J., Cronan J.E. Jr.;
 RT "Growth rate regulation of Escherichia coli acetyl coenzyme A
 carboxylase, which catalyzes the first committed step of lipid
 biosynthesis."
 RL J. Bacteriol. 175:332-340(1993).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: BELONGS TO THE DEDA FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcements/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M68934; AAA23964.1;
 DR EMBL; AE000320; AAC75377.1;
 DR EMBL; D90863; BAA16174.1;
 DR EMBL; AE005463; AAG57446.1;
 DR EMBL; AP002561; BAB36624.1;
 DR EMBL; M15543; AAR24314.1;
 DR EMBL; X02743; CAA26523.1;
 DR EMBL; S53037; AAB24893.1;
 DR PIR; A98029; A98029.
 DR PIR; B29803; XMECAD.
 DR EcoGene; EG10216; deda.
 DR InterPro; IPR00252; Deda.
 DR Pfam; PF00597; Deda; 1.
 KW Transmembrane; Complete proteome.
 FT TRANSMEM 25 45 POTENTIAL.
 FT TRANSMEM 46 66 POTENTIAL.
 FT TRANSMEM 73 93 POTENTIAL.
 FT TRANSMEM 161 181 POTENTIAL.
 FT TRANSMEM 188 208 POTENTIAL.
 SQ SEQUENCE 219 AA; 24510 MW; 37F39ABAB8678C20 CRC64;

Query Match 36.8%; Score 7; DB 1; Length 219;
 Best Local Similarity 100.0%; Pred. No. 15;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 RAARRAA 13
 DB 212 RAARRAA 218

RESULT 13

TRUA_METKA STANDARD; PRT; 256 AA.
 ID TRUA_METKA
 AC Q8TW23;
 DT 15-SEP-2003 (Rel. 42, Created)
 DT 15-SEP-2003 (Rel. 42, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE tRNA pseudouridine synthase A (EC 4.2.1.70) (pseudouridylate synthase
 DE I) (pseudouridine synthase I) (uracil hydrolyase).
 GN TRUA OR MK0888.


```

OS Methanopyrus kandleri.
OC Archaea; Euryarchaeota; Methanopyri; Methanopyrales; Methanopyraceae;
OC Methanopyrus.
OX NCBI_TaxID=2320;
[1]
RN
RP
RC STRAIN-AV19 / DSM 6324 / JCM 9639;
RX MEDLINE=21927647; PubMed=11930014;
RA Slesarev A.I., Mezheva K.V., Makarova K.S., Polushin N.N.,
RA Shcherbinina O.V., Shakhova V.V., Belova G.I., Aravind L.,
RA Natale D.A., Rogozin I.B., Tatusov R.L., Wolf Y.I., Stetter K.O.,
RA Malykh A.G., Koonin E.V., Kozayavkin S.A.;
RT "The complete genome of hyperthermophile Methanopyrus kandleri AV19
RT and monophyly of archaeal methanogens.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:4644-4649(2002).
CC -1- FUNCTION: Formation of pseudouridine at positions 38, 39 and 40 in
CC the anticodon stem and loop of transfer RNAs (By similarity).
CC -1- CATALYTIC ACTIVITY: Uracil + D-ribose 5-phosphate - pseudouridine
CC 5'-phosphate + H(2)O.
CC -1- SIMILARITY: Belongs to the pseudouridine synthase trua family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AE010378; AM02101.1;
CC HAVAP; MF_00171;
CC InterPro: IPR001406; Pseudou_synth_1.
CC Pfam: PF01416; Pseudou_synth_1; 2.
CC Lyase; TRNA processing; Complete proteome.
KW ACT SITE 49 BY SIMILARITY.
FT ACT SITE 49
SQ SEQUENCE 256 AA; 29254 MW; 136C190E515968A4 CRC64;

Query Match 36.8%; Score 7; DB 1; Length 256;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 AAAAAAR 14
DB 116 AAAAAAR 122
|||||

RESULT 14
YOC5_MYCLE
ID YOC5_MYCLE STANDARD; PRT; 323 AA.
AC 069538;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein ML0213.
GN ML0213 OR MLCB2548.18C.
OS Mycobacterium leprae.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1769;
[1]
RN
RP
RC STRAIN-TN;
RC SEQUENCE FROM N.A.
RC MEDLINE=21128732; PubMed=11234002;
RA Cole S.T., Eiglmeier K., Parkhill J., James K.D., Thomson N.R.,
RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,
RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,
RA Holroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,
RA Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA Barrall B.G.;
RT "Massive gene decay in the leprosy bacillus.";

```

```

RL Nature 409:1007-1011(2001).
CC -1- SIMILARITY: BELONGS TO THE UPF0072 (MESJ/YCF62) FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AL023093; CA118805.1;
CC EMBL: AL583917; CAC29721.1;
CC PIR: E86935; E86935.
CC Leptoma; ML0213.
CC InterPro: IPR000541; UPF0021.
CC Pfam: PF01171; ATP_bind3; 1.
CC Hypothetical protein; Complete proteome.
KW SEQUENCE 323 AA; 34113 MW; CA5E345BC863F936 CRC64;
SQ
Query Match 36.8%; Score 7; DB 1; Length 323;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 AAAAAAR 14
DB 102 AAAAAAR 108
|||||

RESULT 15
Y083_CAUCR
ID Y083_CAUCR STANDARD; PRT; 326 AA.
AC P37895;
DT 01-OCT-1994 (Rel. 30, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Hypothetical protein CC2483.
GN CC2483.
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacterales;
OC Caulobacteraceae; Caulobacter.
OX NCBI_TaxID=155892;
[1]
RN
RP
RC STRAIN-ATCC 19089 / CB15;
RC STRAIN-ATCC 19089 / CB15;
RX MEDLINE=93133840; PubMed=8421698;
RA Wang S.P., Sharma P.L., Schoenlein P.V., Ely B.;
RT "A histidine protein kinase is involved in polar organelle
RT development in Caulobacter crescentus.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:630-634(1993).
RN [2]
RP
RC SEQUENCE FROM N.A.
RC STRAIN-ATCC 19089 / CB15;
RX MEDLINE=21173698; PubMed=11259647;
RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA Deboy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
RA Uitterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
RT "Complete genome sequence of Caulobacter crescentus.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
CC -1- SIMILARITY: BELONGS TO THE ARKX FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC

```

DR EMBL; M91449; AAA23053.1; -.
DR EMBL; AE005917; AAK24454.1; -.
DR PIR; B87557; B87557.
DR PIR; S27534; S27534.
DR TIGR; CC2483; -.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR005129; ArgK.
DR Pfam; PF03308; ArgK; 1.
DR SMART; SM00382; AAA; 1.
DR TIGRFAMs; TIGR00750; lso; 1.
KW Hypothetical protein; ATP-binding; Complete proteome.
FT NP_BIND 61 68 ATP (POTENTIAL).
FT CONFLICT 77 77 N -> K (IN REF. 1).
SQ SEQUENCE 326 AA; 34550 MW; 3F58765201CEF32D CRC64;

Query Match 36.8%; Score 7; DB 1; Length 326;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 10 ARAARRA 16
Db 264 ARAARRA 270
|||||

Search completed: August 9, 2003, 16:29:51
Job time : 10.7714 secs